

УДК 681.3

В.А. Алексєєв, В.В. Мостовий, В.С. Терещенко

**ДЕЯКІ МОРФОЛОГІЧНІ АСПЕКТИ АВТОМАТИЗОВАНОГО
ФОРМУВАННЯ ВІДМІНКОВИХ ФОРМ В ДОКУМЕНТАХ
ІНФОРМАЦІЙНИХ СИСТЕМ**

Наводяться шляхи застосування програмних засобів, за допомогою яких проводиться відмінювання ідентифікаційних реквізитів осіб в базі даних за відмінками при формуванні відповідно до запиту, вихідного документа інформаційної системи залежно від потреби текстового змісту або при створенні всього набору відмінків реквізитів для їх зберігання в БД.

Вступ

У зв'язку з широким використанням ідентифікаційних реквізитів осіб у деяких інформаційних системах інколи виникають проблеми стосовно їх відмінювання за відмінками за потребою тексту того чи іншого вихідного документа, в якому знаходяться ці реквізити. Це стосується систем обліку працівників підприємств та установ, пацієнтів лікувальних закладів, відпочиваючих в оздоровчих закладах, клієнтів банківських структур, туристів туристичних агентств, військово-службовців, членів політичних партій, мешканців будинків у житлово-експлуатаційних конторах, виборців у виборчих округах або громадян держави під час перепису населення тощо. Одним словом, в тих системах, де у якості атрибутів використовуються ідентифікаційні реквізити осіб (прізвище, ім'я та по батькові). Особливо часто ці проблеми (необхідність застосування відмінкових форм названих реквізитів) виникають в кадрових системах, наприклад при формуванні наказів у зв'язку із заохоченням працівника, присвоєнням йому чергового звання, нагородженням тощо. Для вирішення такої проблеми існують декілька шляхів.

1. Можливі шляхи створення**відмінкових форм**

Перший шлях полягає в тому, що в багатьох інформаційних системах у базі даних створюються надлишкові поля, де зберігаються реквізити у необхідних відмінках. Це, по-перше, збільшує обсяг бази даних з усіма впливаючими з того наслідками, а, по-друге, збільшує як навантаження на осіб, що збирають та вводять у БД необхідну інформацію, так і час на проведення цих операцій. А ще треба мати на увазі складність самої операції збору такої інформації. Уявіть собі вираз обличчя людини, з якої "знімають" таку інформацію. "Скажіть, а як написати ваше прізвище у родовому відмінку? А як — у давальному...? А як ім'я...? А як...?". І все це відбувається у черзі, скажімо, в стройовій частині або пункті пропуску через державний кордон. Скільки для цього потрібно часу, емоційного та нервового напруження як працівників служб збору інформації, так і тих людей, що стоять у чергах та відповідають на такі запитання. Є ще третій недолік такого підходу: збільшується інформаційний потік між клієнтською та серверною сторонами, якщо інформаційна система побудована за архітектурою "клієнт-сервер" та ще й є територіально розподіленою, корпоративною.

Відповідно до другого шляху подолання цієї проблеми БД також повинна мати надлишкові поля, як і в попередньому випадку, але реквізити у відмінках вносяться у БД оператором у людино-машинному режимі тільки при її модифікації, що впорядковує роботу служби збору інформації та значно знижує емоційне навантаження на неї. Відмінки ж класифікаційних реквізитів (наприклад: військові звання, посади, вчені ступені та звання тощо) можуть вводиться у БД як при створенні, так і при модифікації класифікатора.

Згідно третього шляху, найбільш цікавого з точки зору програмістських уподобань, відбувається створення відмінкових форм реквізитів програмно, у повній відповідності до граматичних правил відмінювання слів по відмінках, під час формування вихідного документа, а не в людино-машинному режимі, як це відбувалось у попередніх шляхах. При цьому зменшуються обсяги бази даних, робіт при зборі інформації та обслуговуванні БД, а також інформаційних потоків між клієнтською та серверною сторонами, але виникає потреба у спеціальному програмному забезпеченні – програмному додатку з відмінювання по відмінках слів-реквізитів відповідно до морфології (розділу мовознавства, що вивчає будову слова залежно від морфем, з яких воно складається, форми словозміни – відмінювання слів за числами, відмінками, родами, особами та ін. [1]) української мови.

Вказаний програмний додаток за трудомісткістю його розробки може бути (на наш погляд) двох видів залежно від підходу:

- "табличний" (третій шлях), що дозволяє проводити відмінювання за відмінками за допомогою спеціальних таблиць відмінкових закінчень (три таблиці для прі-

звища, імені та по батькові для чоловічих реквізитів і три відповідні таблиці для жіночих реквізитів);

- "граматичний" (четвертий шлях), який у повній мірі враховує всі морфологічні правила української граматики з відмінювання прізвищ, імен та по батькові за відмінками разом з існуючими виключеннями з тих правил.

На рис. 1 наведено композиційну граф-схему можливих шляхів та процедур створення відмінкових форм ідентифікаційних реквізитів для формування вихідних документів інформаційної системи.

Розглянемо детальніше два шляхи (за прийнятою в цій статті нумерацією – третій і четвертий), що потребують розробки програмних засобів створення відмінкових форм ідентифікаційних реквізитів для формування вихідних документів інформаційної системи.

2. "Табличний" підхід до створення відмінкових форм

Реалізація і впровадження "табличного" (третього) шляху потребує мінімального часу – 2-3 людино-дні для програміста середньої кваліфікації. Для цього потрібні лише вказані таблиці (таблиці групи "ч" – табл. 1, 2, 3 і таблиці групи "ж" – табл. 4, 5, 6) та досвід розробок систем такого класу. Таблиці наведені як приклад, на якому розглядаються підходи до шляхів розв'язання поставленої проблеми для апробації їх на широкому загалі. Тому їхні обсяги можуть бути збільшені за рахунок проведення додаткових досліджень у разі виникнення такої потреби.

Вирішення задачі побудови відмінкових форм ідентифікаційних реквізитів потребує структурування закінчень слів на загальному полі відмінкових

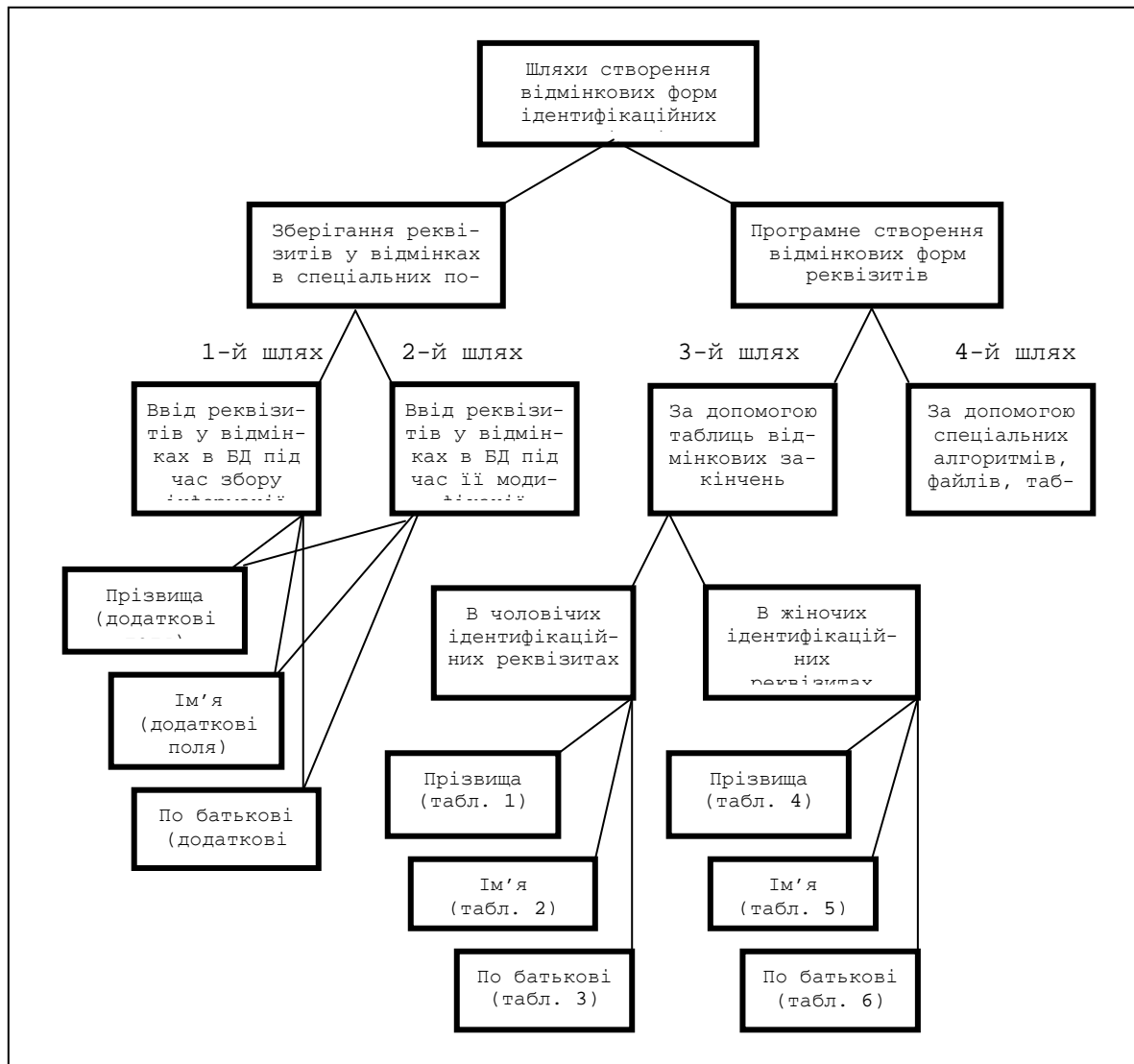


Рис. 1. Композиційна граф-схема можливих шляхів побудови відмінкових форм

закінчень (ВЗ) відповідно до граматичних правил в так званий чотиривимірний простір відмінкових закінчень (ПВЗ) у вигляді таблиць БД. Вони, на перший погляд, нагадують таблиці, що наводяться у підручниках української мови [2], але разом з тим суттєво відрізняються. Відомо, що відмінкові закінчення реквізиту у тому чи іншому відмінку залежать від його закінчення у називному відмінку, відміни, групи та роду. Але при цьому мінімізованому підході відміни та групи визначатися не будуть для зменшення працевкладень: це більш притаманне четвертому підходу. Тому головними таблицями у БД з цього питання будуть таблиці відповідності відмінкових за-

кінчень для чоловічих та жіночих (тут мається на увазі не рід, а стать) реквізитів (перший вимір ПВЗ) для прізвищ, імен та по батькові (другий вимір ПВЗ) таким закінченням реквізиту у називному відмінку (РНВ). Треба ще зауважити, що в нашому випадку закінчення по відмінках визначатиметься дещо інакше, ніж у підручниках. На те є суто програмно-технічні причини. Будемо вважати, що відмінковим закінченням для РНВ є три останні літери. Саме ці останні три літери РНВ (третій вимір ПВЗ) і аналізуватимуться: спочатку остання, потім – передостання, потім передпередостання і по результатах такого аналізу визначатиметься (з вище згаданих таблиць)

дійсне відмінкове закінчення для цього реквізиту у будь-якому іншому заданому відмінку (четвертий вимір ПВЗ). Тут свідомо ми відійшли від поділу РНВ на морфеми (мовознавча значення частини слова – корінь, префікс, суфікс, закінчення [1]), щоб не виникло необхідності в їхньому аналізуванні, що, зрозуміло, потягне за собою низку непростих у вирішенні проблем граматичного характеру (визначення кореню через пошук однокорінних слів, суфіксів, закінчень, для чого необхідні відповідні БД та програмні засоби, не кажучи вже про алгоритми процесів аналізу), як це буде видно при розгляданні іншого ("граматичного") підходу.

Позначимо останню літеру реквізиту у називному відмінку через b_1 , де b – означає наявність літери, а індекс – місце її у реквізиті починаючи з останнього – перша від кінця, b_2 – друга, b_3 – третя від кінця і т.д. Тоді у загальному випадку будь-який n -літерний реквізит у

називному відмінку формально можна записати як множину $B_{n,1}$ або як впорядковану сукупність n літер:

$$B_{n,1} \Leftrightarrow (b_n, \dots, b_i, \dots, b_1), \text{ де } i \in \{n, \dots, 2, 1\}. \quad (1)$$

Визначення індексів – місць літер у впорядкованій сукупності – відбувається тільки для РНВ. Вказана множина літер для РНВ складається з двох не перерізаючих одна одну підмножин: підмножини $B_{n,i}$ літер незмінної основи (індекс i визначає місце останньої її літери) РНВ і підмножини $B_{i-1,1}$ літер відмінкового закінчення для РНВ. Зважаючи на домовленість про трилітерний розмір ВЗ для РНВ можна записати $B_{3,1} \Leftrightarrow (b_3, b_2, b_1)$. Тобто для аналізування РНВ треба розглядати його як теоретико-множинне об'єднання двох підмножин

Таблиця 1. Відмінкові закінчення для чоловічого ідентифікаційного реквізиту – прізвище

№	Останні літери реквізиту у називному відмінку			Формальний запис реквізиту по відмінках					
	$[b_3]$	$[b_2]$	$[b_1]$	Р	Д	З	О	М	К
1	Будь-яка	Приг. (крім ж, ч, ш, щ)	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,2}+і$	
2	Будь-яка	г	а	$B_{n,2}+и$	$B_{n,3}+зі$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+зі$	
3	Будь-яка	к	а	$B_{n,2}+и$	$B_{n,3}+ці$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+ці$	
4	Будь-яка	х	а	$B_{n,2}+и$	$B_{n,3}+сі$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+сі$	
5	Будь-яка	ж, ч, ш, щ	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+єю$	$B_{n,2}+і$	
6	Будь-яка	Голосна (крім и)	й	$B_{n,2}+ю$	$B_{n,2}+ю$	$B_{n,1}$	$B_{n,2}+єм$	$B_{n,2}+ї$	
7	Будь-яка	и	й	$B_{n,3}+ого$	$B_{n,3}+ому$	$B_{n,3}+ого$	$B_{n,3}+им$	$B_{n,3}+ому$	
8	Будь-яка	Приголос-на	е	$B_{n,2}+я$	$B_{n,2}+ю$	$B_{n,2}+е$	$B_{n,2}+єм$	$B_{n,2}+і$	
9	Будь-яка	ж, ч, ш, щ	е	$B_{n,2}+а$	$B_{n,2}+у$	$B_{n,2}+е$	$B_{n,2}+єм$	$B_{n,2}+і$	
10	Будь-яка	Приголос-на	о	$B_{n,2}+а$	$B_{n,2}+у$	$B_{n,2}+а$	$B_{n,2}+ом$	$B_{n,2}+ові$	
11	Будь-яка	Приголос-на	я	$B_{n,2}+і$	$B_{n,2}+і$	$B_{n,2}+ю$	$B_{n,2}+єю$	$B_{n,2}+і$	
12	Будь-яка	Будь-яка	Приголос-на	$B_{n,1}+а$	$B_{n,1}+у$	$B_{n,1}$	$B_{n,1}+ом$	$B_{n,1}+і$	
13	Будь-яка	Будь-яка	ж, ч, ш, щ	$B_{n,1}+а$	$B_{n,1}+еві$	$B_{n,1}+а$	$B_{n,1}+єм$	$B_{n,1}+еві$	
14	Будь-	Приголос-	ь	$B_{n,1}+я$	$B_{n,1}+еві$	$B_{n,1}+я$	$B_{n,1}+єм$	$B_{n,1}+еві$	

$$B_{n,1} \Leftrightarrow B_{n,4} + (b_3, b_2, b_1). \quad (2)$$

Для інших відмінків такого запису зробити не можна тому, що розміри відмінкових закінчень у цьому разі можуть бути будь-якими – від нуля до чотирьох. Тому тут буде справедливим інший формальний запис:

$$B_{n,1} \Leftrightarrow B_{n,i} + (b_x, \dots, b_1), \quad (3)$$

де $B_{n,i}$ – підмножина літер незмінної основи, яка визначається на множині літер для РНВ;

(b_x, \dots, b_1) – підмножина літер відмінкового закінчення ($x \in \{0, \dots, m\}$, а m на практиці не перевищує 4). Зрозуміло, якщо $x = 0$, то підмножина (b_x, \dots, b_1) вироджується – закінчення немає

і, значить, відмінкова форма реквізиту у цьому відмінку складається тільки з незмінної основи.

Ці підмножини вказані в таблицях у графах всіх інших відмінків у вигляді запису через знак "+" теоретико-множинного об'єднання двох підмножин, причому перша підмножина вказана у формальному вигляді, а друга – у вигляді впорядкованої сукупності. Треба зауважити, що так звана незмінна основа для одного й того ж значення реквізиту у різних відмінках може бути різною, тобто значення індексу i у підмножині $B_{n,i}$ для різних відмінків може не співпадати. Термін "відмінкове закінчення", що використовується в цій статті, не є морфемою, а є певною впорядкованою сукупністю літер, яка може

Таблиця 2. Відмінкові закінчення для чоловічого ідентифікаційного реквізиту – ім'я

№	Останні літери реквізиту у називному відмінку			Формальний запис реквізиту по відмінках					
	$[b_3]$	$[b_2]$	$[b_1]$	Р	Д	З	О	М	К
1	Будь-яка	Приг. (крі м ж, ч, ш, щ)	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,2}+і$	$B_{n,2}+о$
2	Будь-яка	г	а	$B_{n,2}+и$	$B_{n,3}+зі$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+зі$	$B_{n,2}+о$
3	Будь-яка	к	а	$B_{n,2}+и$	$B_{n,3}+ці$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+ці$	$B_{n,2}+о$
4	Будь-яка	х	а	$B_{n,2}+и$	$B_{n,3}+сі$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+сі$	$B_{n,2}+о$
5	Будь-яка	ж, ч, ш, щ	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+ею$	$B_{n,2}+і$	$B_{n,2}+е$
6	Будь-яка	ж, ч, ш, щ	е	$B_{n,2}+а$	$B_{n,2}+у$	$B_{n,2}+е$	$B_{n,2}+ем$	$B_{n,2}+і$	$B_{n,2}+е$
7	Будь-яка	Голосна	й	$B_{n,2}+я$	$B_{n,2}+ю$	$B_{n,1}$	$B_{n,2}+єм$	$B_{n,2}+і$	$B_{n,2}+ю$
8	Будь-яка	Приголосна	о	$B_{n,2}+а$	$B_{n,2}+у$	$B_{n,2}+а$	$B_{n,2}+ом$	$B_{n,2}+ові$	$B_{n,2}+е$
9	Будь-яка	Приголосна	я	$B_{n,2}+і$	$B_{n,2}+і$	$B_{n,2}+ю$	$B_{n,2}+ею$	$B_{n,2}+і$	$B_{n,2}+ю$
10	Будь-яка	Будь-яка	Приголосна	$B_{n,1}+а$	$B_{n,1}+у$	$B_{n,1}$	$B_{n,1}+ом$	$B_{n,1}+і$	$B_{n,1}+е$

співпадати з морфемою "закінчення". Крім того, кличний відмінок у цих таблицях застосовується лише для імен та по батькові, а перед прізвищем в такому разі у кличному відмінку ставиться лише звернення (наприклад, пане, товаришу, громадянине або пані, панно, товаришко, громадянко тощо).

В таблицях прийняті такі умовні позначення:

Р, Д, З, О, М, К – відповідно родовий, давальний, знахідний, орудний, місцевий, кличний відмінки.

Все поле відмінкових закін-

чень (множина M), наведених в цих таблицях, складається з підмножин ВЗ M_r ($r \in \{1,2\}$) чоловічих та жіночих ідентифікаційних реквізитів, які складаються з підмножин ВЗ M_{rv} ($v \in \{1,2,3\}$) ідентифікаційних реквізитів по їх видах, які в свою чергу складаються з підмножин ВЗ M_{rvz} ($z \in \{1,...,Z_r\}$) з однаковими закінченнями РНВ, складених з підмножин ВЗ M_{rvzw} ($w \in \{1,...,6\}$) всього набору конкретних відмінків для даного закінчення РНВ. Такий вербальний опис співвідношення на-

Таблиця 3. Відмінкові закінчення для чоловічого ідентифікаційного реквізиту – по батькові

№	Останні літери реквізиту у називному відмінку			Формальний запис реквізиту по відмінках					
	$[b_3]$	$[b_2]$	$[b_1]$	Р	Д	З	О	М	К
1	Будь-яка	Голосна	ч	$B_{n,1}+а$	$B_{n,1}+у$	$B_{n,1}+а$	$B_{n,1}+ем$	$B_{n,1}+і$	$B_{n,1}+у$

Таблиця 4. Відмінкові закінчення для жіночого ідентифікаційного реквізиту – прізвище

№	Останні літери реквізиту у називному відмінку			Формальний запис реквізиту по відмінках					
	$[b_3]$	$[b_2]$	$[b_1]$	Р	Д	З	О	М	К
1	Будь-яка	Приг. (крім ж, ч, ш, щ)	а	$B_{n,2}+o\ddot{i}$	$B_{n,2}+i\ddot{y}$	$B_{n,2}+y$	$B_{n,2}+o\ddot{y}$	$B_{n,2}+i\ddot{y}$	
2	Будь-яка	г	а	$B_{n,2}+и$	$B_{n,3}+z\ddot{i}$	$B_{n,2}+y$	$B_{n,2}+o\ddot{y}$	$B_{n,3}+z\ddot{i}$	
3	Будь-яка	к	а	$B_{n,2}+и$	$B_{n,3}+ц\ddot{i}$	$B_{n,2}+y$	$B_{n,2}+o\ddot{y}$	$B_{n,3}+ц\ddot{i}$	
4	Будь-яка	х	а	$B_{n,2}+и$	$B_{n,3}+c\ddot{i}$	$B_{n,2}+y$	$B_{n,2}+o\ddot{y}$	$B_{n,3}+c\ddot{i}$	
5	Будь-яка	ж, ч, ш, щ	а	$B_{n,2}+i$	$B_{n,2}+i$	$B_{n,2}+y$	$B_{n,2}+e\ddot{y}$	$B_{n,2}+i$	
6	Будь-яка	Приголос-на	е	$B_{n,2}+и$	$B_{n,2}+i$	$B_{n,2}+e$	$B_{n,2}+e\ddot{y}$	$B_{n,2}+i$	
7	Будь-яка	ж, ч, ш, щ	е	$B_{n,2}+e$	$B_{n,2}+e$	$B_{n,2}+e$	$B_{n,2}+e$	$B_{n,2}+e$	
8	Будь-яка	Голосна (крім и)	й	$B_{n,2}+i\ddot{}$	$B_{n,2}+i\ddot{}$	$B_{n,2}+y\ddot{}$	$B_{n,2}+e\ddot{y}$	$B_{n,2}+i\ddot{}$	
9	Будь-яка	и	й	$B_{n,3}+и\ddot{y}$	$B_{n,3}+и\ddot{y}$	$B_{n,3}+и\ddot{y}$	$B_{n,3}+и\ddot{y}$	$B_{n,3}+и\ddot{y}$	
10	Будь-яка	Приголос-на	о	$B_{n,2}+o$	$B_{n,2}+o$	$B_{n,2}+o$	$B_{n,2}+o$	$B_{n,2}+o$	
11	Будь-яка	Приголос-на	я	$B_{n,2}+i$	$B_{n,2}+i$	$B_{n,2}+y\ddot{}$	$B_{n,2}+e\ddot{y}$	$B_{n,2}+i$	
12	Будь-яка	Будь-яка	Приголос-на	$B_{n,1}$	$B_{n,1}$	$B_{n,1}$	$B_{n,1}$	$B_{n,1}$	
13	Будь-яка	Приголос-на (крім	ь	$B_{n,2}+i$	$B_{n,2}+i$	$B_{n,2}+y\ddot{}$	$B_{n,2}+e\ddot{y}$	$B_{n,2}+i$	

званих підмножин формально можна записати у вигляді

$$M = \sum_r M_r = \sum_r \sum_v M_{rv} = \sum_r \sum_v \sum_z M_{rvz} = \sum_r \sum_v \sum_z \sum_w M_{rvzw}, \quad (4)$$

де r – індекс роду (статі володаря) реквізиту;

v – індекс виду реквізиту;

z – індекс номера рядка в таблиці;

w – індекс необхідного відмінку.

Ці взаємно вкладені підмножини ВЗ і створюють вже визначений раніше простір відмінкових закінчень – ПВЗ, який має чотири

виміри і схематично зображений на рис. 2.

Для отримання відмінкової форми реквізиту необхідно у цьому чотиривимірному просторі відмінкових закінчень визначити:

- підмножину ВЗ – групу таблиць (табл.1, 2, 3 або табл.4, 5, 6) – залежно від першого виміру ПВЗ – роду r реквізиту: чоловічий ($r=1$), жіночий ($r=2$):

$$M_r = \begin{cases} \text{табл.}(1, 2, 3) & \text{при } r=1, \\ \text{табл.}(4, 5, 6) & \text{при } r=2; \end{cases}$$

- підмножину ВЗ – номер таблиці (табл. 1, 4 або табл. 2, 5, або табл. 3, 6) – залежно від друго-

Таблиця 5. Відмінкові закінчення для жіночого ідентифікаційного реквізиту – ім'я

№	Останні літери реквізиту у називному відмінку			Формальний запис реквізиту по відмінках					
	$[b_3]$	$[b_2]$	$[b_1]$	Р	Д	З	О	М	К
1	Будь-яка	Приг. (крім ж, ч, ш, щ)	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,2}+і$	$B_{n,2}+о$
2	Будь-яка	г	а	$B_{n,2}+и$	$B_{n,3}+з+і$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+зі$	$B_{n,2}+о$
3	Будь-яка	к	а	$B_{n,2}+и$	$B_{n,3}+ці$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+ці$	$B_{n,2}+о$
4	Будь-яка	х	а	$B_{n,2}+и$	$B_{n,3}+сі$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,3}+сі$	$B_{n,2}+о$
5	Будь-яка	ж, ч, ш, щ	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+ею$	$B_{n,2}+і$	$B_{n,2}+е$
6	Будь-яка	Голосна (крім и)	й	$B_{n,2}+ю$	$B_{n,2}+ю$	$B_{n,1}$	$B_{n,2}+єм$	$B_{n,2}+і$	$B_{n,2}+ю$
7	Будь-яка	Приголосна	е	$B_{n,2}+я$	$B_{n,2}+ю$	$B_{n,2}+е$	$B_{n,2}+єм$	$B_{n,2}+і$	$B_{n,2}+е$
8	Будь-яка	Приголосна	я	$B_{n,2}+і$	$B_{n,2}+і$	$B_{n,2}+ю$	$B_{n,2}+ею$	$B_{n,2}+і$	$B_{n,2}+ю$
9	Будь-яка	Будь-яка	Приголосна	$B_{n,1}+и$	$B_{n,1}+і$	$B_{n,1}$	$B_{n,1}+ою$	$B_{n,1}+і$	$B_{n,1}+е$

го виміру ПВЗ – виду v реквізиту: $r=1$ – прізвище ($v=1$), ім'я ($v=2$), по батькові ($v=3$):

$$M_{r=1,v} = \begin{cases} \text{табл. 1} & \text{при } v=1; \\ \text{табл. 2} & \text{при } v=2; \\ \text{табл. 3} & \text{при } v=3; \end{cases}$$

$$M_{r=2,v} = \begin{cases} \text{табл. 4} & \text{при } v=1; \\ \text{табл. 5} & \text{при } v=2; \\ \text{табл. 6} & \text{при } v=3; \end{cases}$$

• підмножину ВЗ – ВЗ рядка z в таблиці – залежно від третього виміру ПВЗ – результату аналізу трьох останніх літер РНВ (b_3, b_2, b_1) :

$M_{rvz} = f(b_3, b_2, b_1)$ при фіксованих r, v ;

• ВЗ в графі (Р, Д, З, О, М, К) в таблиці залежно від четвертого виміру ПВЗ – необхідного відмінку w : родовий ($w=1$), давальний ($w=2$), знахідний ($w=3$), орудний ($w=4$), місцевий ($w=5$), кличний ($w=6$) відмінки: $M_{rvzw} = f(w)$ при фіксованих r, v, w .

Значення Z_{rv} – кількість рядків ВЗ в таблицях – може бути неоднаковим для різних таблиць. До того ж воно не фіксоване, тому що обсяги таблиць, як вже було сказано, можуть бути збільшені за рахунок проведення додаткових досліджень.

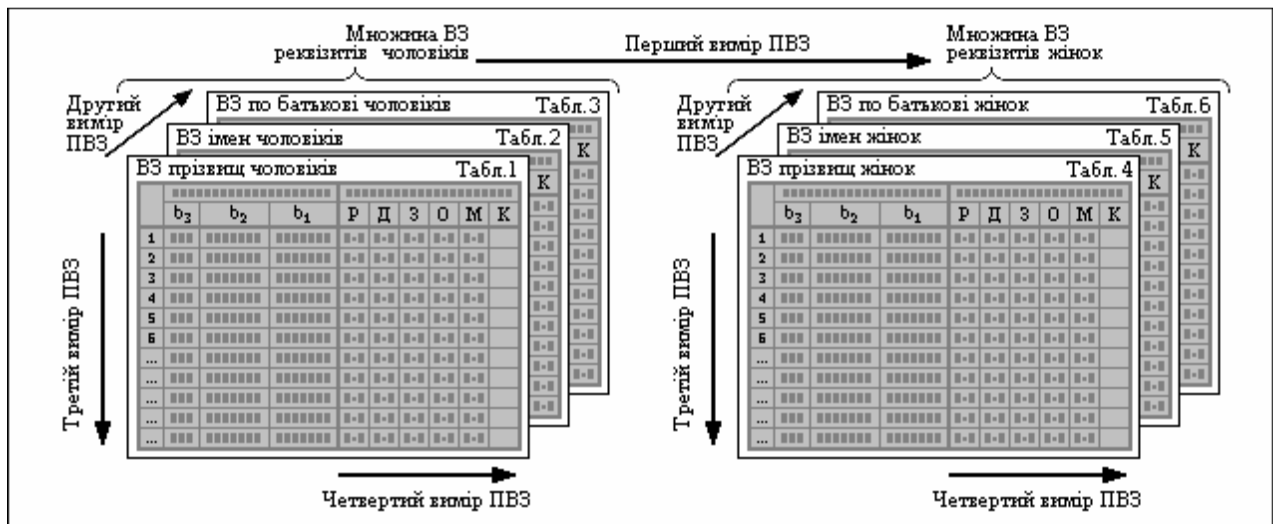


Рис. 2. Чотиривимірний простір відмінкових закінчень

Результатом такого визначення стане формальний запис відмінка заданого реквізиту — $[B_{n,i}+VZ]$, який у свою чергу визначає незмінну основу (підстроку строкової перемінної — РНВ, починаючи з n -ї і закінчуючи i -ю літерою) та відмінкове закінчення. Грунтуючись на цих даних, за допомогою конкатенації будується пошукувана відмінкова форма реквізиту відповідно до алгоритму, наведеному у табличній формі (табл. 7). Приклад: побудувати форму родового відмінку від чоловічого прізвища Таратута відповідно до наведеного вище алгоритму (табл. 8).

Більш наглядно послідовність операцій при виконанні завдання видно на схемі, наведеній на рис. 3.

Блок-схему алгоритму при цьому підході визначення закінчень у тому чи іншому відмінку реквізиту залежно від його закінчення у називному відмінку та роду тут не наводимо через його тривіальність: кожний програміст-початківець може його створити, що теж є однією з позитивних рис такого підходу. А блок-схемним методом користуються при побудові складних алгоритмів, у якості попередньої перевірки і наглядного представлення логіки розв'язання задачі [3, с. 63].

Таблиця 6. Відмінкові закінчення для жіночого ідентифікаційного реквізиту — по батькові

№	Останні літери реквізиту у називному відмінку			Формальний запис реквізиту по відмінках					
	$[b_3]$	$[b_2]$	$[b_1]$	Р	Д	З	О	М	К
1	Будьяка	Приголосна	а	$B_{n,2}+и$	$B_{n,2}+і$	$B_{n,2}+у$	$B_{n,2}+ою$	$B_{n,2}+і$	$B_{n,2}+о$

Таблиця 7. Алгоритм побудови відмінкових форм реквізитів

№	Операції	Умови проведення операцій	Результати операцій
		Знайти для заданого рек	Вид реквізиту

Таблиця 8. Приклад побудови форми родового відмінка від чоловічого прізвища Таратута відповідно до наведеного вище алгоритму

№	Операції	Умови проведення операцій	Результати операцій
1	Визначення умов за- вдання	Знайти родовий відмінок чоловічого прізвища "Та- ратута"	Реквізит – Таратута Вид реквізиту – прізвище Рід реквізиту – чоловічий Необхідний відмінок – ро- довий
2	Визначення групи таб- лиць	Рід реквізиту – чолові- чий	Група таблиць "ч" (1, 2, 3)
3	Визначення номеру таб- лиці в групі	Вид реквізиту – прізвище	Таблиця 1
4	Визначення рядка таб- лиці по останніх трьох літерах	$b_3 \Leftrightarrow$ будь-яка $b_2 \Leftrightarrow$ приголосна, $b_1 \Leftrightarrow$ а	Рядок 1
5	Визначення графи таб- лиці для заданого від- мінка	Родовий відмінок	Графа "Р"
6	Визначення формального запису реквізиту у по- трібному відмінку	Табл.1, рядок 1, графа "Р"	$B_{n,2}+и$
7	Побудова для потрібно- го відмінка незмінної основи реквізиту (на- приклад, функція SUBSTR для VFP 3.0)	Незмінна основа від РНВ для родового відмінка – $B_{n,2}$	Таратут

3. "Граматичний" підхід до ство- рення відмінкових форм

Застосування "граматичного" (четвертого) підходу до розробки програмних засобів створення відмінкових форм ідентифікаційних реквізитів для формування вихідних документів інформаційної системи значно більш трудомісткий. Для його реалізації і впровадження розробленого для цього програмного засобу потрібно витратити набагато більше часу кваліфікованих програмістів, системних аналітиків та фахівців з української морфології. При виборі такого підходу до подолання проблеми відмінювання ідентифікаційних реквізитів виникає потреба в проведенні додаткових досліджень, пов'язаних зі створенням спеціальних файлів вживаних іменників, які могли б при словотворенні прізвищ лягти в їх основу, наприклад перелік імен-

ників (і не тільки), пов'язаних з професією, з народним побутом, вуличними прізвиськами, іменами, національністю тощо, всім тим, чим займається ономастика (ономастологія) як розділ загальної лексикології [4]. Крім того, необхідно ретельно врахувати вимоги морфологічних правил. При цьому виникає ряд ускладнень:

- прізвища, як іменники, можуть бути тільки чоловічого або жіночого роду відповідно до статі володаря, але зустрічаються вони і з усіма ознаками іменника середнього роду другої відміни, тому що мають закінчення -о, -е (наприклад: Терещенко, Собко, Власе), а ще зустрічаються – подвійного роду (той же приклад);

- є такі прізвища, що мають властивості прикметників і, будучи субстивованими назвами, відмінюються як прикметники тве-

рдої чи м'якої групи, але якщо ж в їх основі є суфікси -ів, -їв, -ов, -ова, -ев, -ева, -ин, -ін, -їн, -ськ, то вони відмінюються як іменники (наприклад: Мостовий, Коцюбинський);

- велика кількість чоловічих та жіночих прізвищ співпада-

тира, Ротар – Ротаря);

- прізвища, що мають форму іменників II відміни чоловічого роду з кінцевими -яр (наприклад: Котляр) можуть належати і до твердої групи і до мішаної, у зв'язку з чим мають різні закінчення в однакових, крім називного, відмінках залежно від місця

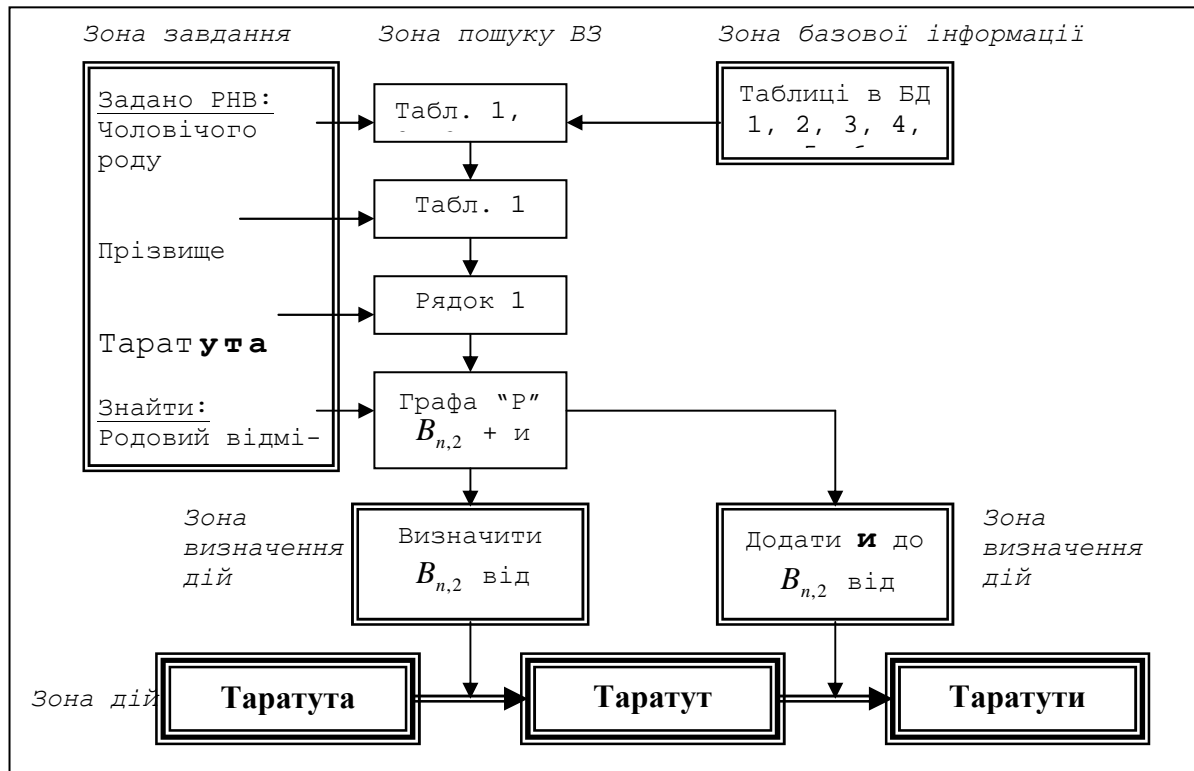


Рис. 3. Схема операцій виконання завдання з знаходження родового відмінка прізвища

ють (приклад першого пункту) а закінчення в однакових відмінках – різні;

- на відміну від РНВ в коренях прізвищ у давальному та місцевому відмінках має місце чергування приголосних літер к-ц, г-з, х-с (наприклад: Притика – Притиці, Кочерга – Кочерзі, Макуха – Макусі);

- прізвища, що мають форму іменників II відміни чоловічого роду з кінцевими -ир, -ар можуть належати і до твердої групи і до м'якої, у зв'язку з чим мають різні закінчення в однакових, крім називного, відмінках залежно від місця знаходження наголосу (наприклад: Волонтир – Волон-

знаходження наголосу: на основі чи на закінченні;

- у подвійних прізвищах (наприклад: Гулак-Артемівський, Миклухо-Маклай) треба відмінювати по відмінках обидві частини прізвища;

- іншомовні прізвища більшістю є незмінними, а саме з закінченнями на -их, -ово, -аго, -е, -є, -і, -ї, -о, -у, та -а з попереднім голосним;

- треба враховувати випадіння літер в основі прізвищ при їх відмінюванні (наприклад: Вересень – Вересня);

- відбувається зміна літер в основі прізвищ при їх відміню-

ванні (наприклад: Тхір – Тхора);

- при зовні однакових формах РНВ мають місце різні закінчення та зміни в основі реквізиту в однакових, крім називного, відмінках (наприклад: Звір – Звіра, Тхір – Тхора, Твір – Твору, Горобець – Горобця, Стрілець – Стрільця).

Далеко не всі ці зауваження можна зняти за допомогою наведених вище таблиць. Навіть при поверхневому знайомстві з цими зауваженнями вимушені зробити висновок про необхідність застосування цілої низки відповідних алгоритмів та спеціально створених файлів:

- о алгоритм розпізнавання частини мови, до яких належить прізвище (з ім'ям та по батькові проблем менше);

- о алгоритм розпізнавання роду, до якого належить те чи інше прізвище, тому що відмінкові закінчення в них можуть бути різними залежно від статі власника цього прізвища;

- о алгоритм визначення групи іменників, до яких відноситься

реквізит;

- о алгоритм визначення відмінки іменників, до яких відноситься реквізит;

- о алгоритм розпізнавання основ в прізвищах та відрізнень їх від закінчень;

- о алгоритм розпізнавання морфем (коренів, словотворчих та формотворчих суфіксів, закінчень);

- о алгоритм визначення місця знаходження наголосу – на основі чи на закінченні в прізвищах;

- о файли з іменниками з кінцевими -ир, -ар для твердої і м'якої груп окремо;

- о файли з іменниками з кінцевими -яр для твердої і мішаної груп окремо або алгоритм розпізнавання прізвищ, які визначають осіб за фахом або родом діяльності;

- о файли з двома полями для подвійних прізвищ, тому що кожен з їх частин треба відмінювати окремо тощо.

Відмінювання ПІБ				
№	Згоден	Прізвище	Ім'я	По батькові
1	<input checked="" type="checkbox"/>	АГАДЖАНОВ	Григорій	Сергійович
		Родовий АГАДЖАНОВА	Григорія	Сергійовича
		Давальний АГАДЖАНОВУ	Григорію	Сергійовичу
2	<input type="checkbox"/>	ШИЛО	Борис	Вікторович
		Родовий ШИЛА	Бориса	Вікторовича
		Давальний ШИЛУ	Борису	Вікторовичу
3	<input type="checkbox"/>	ШАМА	Ігор	Володимирович
		Родовий ШАМИ	Ігора	Володимировича
		Давальний ШАМІ	Ігорю	Володимировичу
4	<input type="checkbox"/>	АНАНЬІН	Олег	Валерійович
		Родовий АНАНЬІНА	Олега	Валерійовича
		Давальний АНАНЬІНУ	Олегу	Валерійовичу
5	<input type="checkbox"/>	БАЛКОВИЙ	Юрій	Миколайович
		Родовий БАЛКОВОГО	Юрія	Миколайовича
		Давальний БАЛКОВОМУ	Юрію	Миколайовичу

Показати	<input checked="" type="checkbox"/> Згоден	АГАДЖАНОВ	Григорій	Сергійович
<input checked="" type="radio"/> Всі		Родовий: АГАДЖАНОВА	Григорія	Сергійовича
<input type="radio"/> Невідмічен		Давальний: АГАДЖАНОВУ	Григорію	Сергійовичу

Рис. 4. Екранна форма для відмінювання ідентифікаційних реквізитів ПІБ

Мабуть, нема потреби ще раз наголошувати, що цей шлях дуже трудомісткий. І навряд чи він є оптимальним при створенні всього лише кадрової системи, де обсяг БД не буде перевищувати кілька тисяч записів. У зв'язку з цим виникає запитання: "так який шлях вибрати?" Може, дійсно введення необхідних даних у людино-машинному режимі є оптимальним варіантом. А, може, краще створити файл-довідник прізвищ, що зустрічаються в Україні, перейнявши їх з телефонних книг великих міст нашої держави та створивши їхні відмінкові форми за допомогою фахівців з української морфології. Але, на наш погляд, найбільш прийнятний шлях, як завжди, лежить посередині.

4. Інтегральна технологія створення відмінкових форм

Треба реалізувати інтегральну технологію, в якій будуть застосовані і програмні засоби і людино-машинний режим. Ця технологія найбільш прийнятна для кадрової системи і передбачає реалізацію її у три етапи:

1) завантаження у людино-машинному режимі бази даних відповідними реквізитами у називному відмінку на всіх осіб;

2) створення за допомогою програмних засобів ("табличного" підходу) у відповідних полях БД всіх необхідних відмінкових форм реквізитів;

3) контролювання у людино-машинному режимі результатів створення відмінкових форм реквізитів та, в разі необхідності, редагування помилково створених відмінкових форм ідентифікаційних реквізитів у БД.

Перший етап цієї технології притаманний для всіх шляхів. Другий етап запозичено з другого шляху (див. рис.1), тому цей

етап технології зберігає деякі його недоліки: наприклад, існування в БД надлишкових полів, але значно заощаджує час на завантаженні БД всіма відмінковими формами за рахунок їх автоматичного створення за допомогою програмних засобів під час модифікації БД. Контролювання (третій етап) та редагування вже завантаженої БД, на наш погляд, буде вимагати не більше 5-10% від загального обсягу БД у межах тієї її частини, що призначена для реквізитів.

Саме ця інтегральна технологія використана при розробці проектних рішень і побудові відповідного програмного модуля "Корпоративної автоматизованої інформаційної системи кадрового забезпечення", метою створення якої є підвищення:

- ефективності обліку і обробки даних, що є в персональних облікових картах службовців;

- повноти накопичення, вірогідності та аналізу даних про переміщення осіб по "дереву" організаційно-штатної структури установи з розрахунку періоду праці;

- якості формування звітів і довідок як по регламентованих, так і по довільних формах;

- оперативності та надійності інформаційного обміну між структурними підрозділами установи з питань кадрового забезпечення, а також з підрозділами міністерств, інших органів центральної виконавчої влади при вирішенні задач кадрового забезпечення.

Для досягнення цієї мети необхідно вирішити велике коло задач, пов'язаних з видачею вихідних документів, що часто містять реквізити осіб у будь-якому відмінку. Саме це й допомагає реалізувати запропонована інтегральна технологія. Вона ґрунту-

ється на використанні програмних засобів створення відмінкових форм ідентифікаційних реквізитів та людино-машинного режиму редагування. Сенс підходу міститься у тому, що в базі даних крім полів для називного відмінку резервуються поля для інших відмінків (у даному разі – родового, давального) для прізвищ, імен та по батькові. Для заповнення та редагування цих додаткових полів в режимі завантаження БД системи створено спеціальну екранну форму відмінювання ідентифікаційних реквізитів (рис.4) для спілкування з цим програмним модулем, який є достатньо автономним, щоб його можна було б легко замінити на іншу більш якісну версію, якщо виникне така потреба, у повній відповідності до відомих моделей життєвого циклу програмних продуктів [5].

На екранній формі передбачена спеціальна клавіша <Відмінювати>, при активізації якої програмний засіб, реалізований за алгоритмом "табличного" підходу, створює потрібні відмінки та заповнює відповідні поля. Після цього адміністратор БД переглядає результати відпрацювання програмного засобу і позначає вірні записи, а невірні виправляє в режимі ручного вводу і теж позначає, за допомогою перемикача <Згоден> про те, що запис перевірено і виправлено. Це робиться для того, щоб позначені записи при наступних запусках автоматичного відмінювання не змінювалися. Таким чином, при вводі нових реквізитів після автоматичного відмінювання необхідно буде перевіряти, а при потребі й редагувати, тільки нові реквізити.

Формування довідок з використанням реквізитів у різних відмінках зводиться до витягу з полів БД вже готових ПІБ у необхідному відмінку. Перевага цього

підходу полягає в тому, що при формуванні довідок використовуються вже перевірені записи реквізитів у відмінках, а якби відмінювати ПІБ у процесі формування довідок, то могли б бути невірними результати, наприклад, для іноземних прізвищ та імен, які неможливо було б виправити. Для цього треба було б ввести ще один режим – режим редагування довідок, а це внесло б додаткові складнощі. До того ж, у людино-машинному режимі перевірку правильності проведеного відмінювання може проконтролювати фахівець з української мови один раз (при створенні, модифікації БД або при проведенні регламентних робіт) і не треба його залучати для перевірки кожної довідки.

Оскільки програмні засоби відмінювання працюють на конкретній вже завантаженій БД, де вже відомі рід (стать власника реквізитів), конкретні ПІБ, то відповідає необхідність в алгоритмах розпізнавання цих даних. Спрощений таким чином алгоритм звівся до створення бібліотеки DLL з

множиною точок входу по дванадцяти функціях (дві для чоловічого та жіночого роду, три для ПІБ та дві для родового та давального відмінків), на вхід яких передається реквізит ПІБ у називному відмінку. Код функції й визначає операцію, яку необхідно реалізувати програмному засобу. Наприклад, M_Name_r говорить, що ця функція поверне чоловіче ім'я у родовому відмінку (див. табл. 9).

Результати цих функцій заносяться у додаткові поля БД: три для ПІБ у родовому відмінку (Fam_r, Name_r, Otch_r) та три для ПІБ у давальному відмінку (Fam_d, Name_d, Otch_d). Для внесення спеціальної позначки, яка свідчитиме, що запис переглянуто і він вірний, створюється ще одне логічне поле (Isfile).

Бібліотека DLL з переліченими вище функціями написана у Delphi та підключена як зовнішня бібліотека до системи, реалізованої у середовищі Power Builder.

Таблиця 9. Функції відмінювання реквізитів ПІБ

№	Код функції	Розшифровка параметрів коду функцій		
		Параметр 1 (рід)	Параметр 2 (ПІБ)	Параметр 3 (відмінок)
1	M_Fam_r	Чоловічий	Прізвище	Родовий
2	M_Fam_d	Чоловічий	Прізвище	Давальний
3	M_Name_r	Чоловічий	Ім'я	Родовий
4	M_Name_d	Чоловічий	Ім'я	Давальний
5	M_Otch_r	Чоловічий	По батькові	Родовий
6	M_Otch_d	Чоловічий	По батькові	Давальний
7	F_Fam_r	Жіночий	Прізвище	Родовий
8	F_Fam_d	Жіночий	Прізвище	Давальний
9	F_Name_r	Жіночий	Ім'я	Родовий
10	F_Name_d	Жіночий	Ім'я	Давальний

Висновки

Запропонована інтегральна технологія побудови відмінкових закінчень ідентифікаційних реквізитів в інформаційних системах. Відповідно до цієї технології застосовуються нескладні (для їх розробки не треба великих працевкладень) програмні засоби для автоматичного створення відмінкових форм реквізитів при модифікації БД і людино-машинний режим для всебічного контролю та, за необхідністю, редагування помилково створених відмінкових форм ідентифікаційних реквізитів у БД інформаційної системи. Така технологія не потребує великих працевитрат та втрат часу ані при створенні програмних засобів, ані при зборі інформаційних даних, ані при веденні і модифікації БД, ані при формуванні вихідних документів з ідентифікаційними реквізитами у потрібних відмінках.

1. Словник іншомовних слів. — К.: Голлова редакція Української радянської енциклопедії АН УРСР, 1974. — 776 с.
2. Пентиліук М.І., Іващенко О.В. Українська мова: Підручник-комплект. — К.: Ленвіт, 2001. — 352 с.
3. Алфєрова З.В. Теория алгоритмов. — М.: Статистика, 1973. — 164 с.
4. Советский энциклопедический словарь. — М.: Изд. "СЭ", 1989. — 1632 с.
5. Алексеев В.А., Терещенко В.С. Развитие спиральной модели жизненного цикла программных систем // Пробл. программирования. — 2003. — №4. — С. 34-42.

Отримано 24.02.04

Про авторів

Алексеев Виктор Анатолійович,
канд. техн. наук, завідувач відділу

Місце роботи автора:

Інститут програмних систем НАН України
м. Київ, пр-т Академіка Глушкова, 40

Тел. (044) 266 4228

E-mail: alecseev@isofts.kiev.ua

Мостовий Валентин Васильович,

канд. техн. наук, ст. наук.
співробітник

Місце роботи автора:

Інститут кібернетики НАН України

м. Київ, пр-т Академіка Глушкова, 42

Тел. (044) 266 6321

E-mail: most@isofts.kiev.ua

Терещенко Валерій Савелійович,

канд. техн. наук, ст. наук.
співробітник

Місце роботи автора:

Інститут програмних систем НАН України

м. Київ, пр-т Академіка Глушкова, 40

Тел. (044) 266 6321

E-mail: terek@isofts.kiev.ua